CLAIMS

What is claimed is:

1. A program storage device accessible by a computer, tangible embodying a program of instructions executable by said computer to perform method steps for protein structure alignment, said methods steps comprising of:

(a) receiving a first protein with N_1 atoms;

(b) receiving a second protein with N_2 atoms;

(c) making an initial alignment of said atoms of said first protein to said atoms of said second protein;

(d) calculating all atomic distances between said coordinates of said atoms of said first protein and said atomic coordinates of said atoms of said second protein;

(e) defining a matrix with a plurality of binary assignment variables wherein each binary assignment variable corresponding to either a match or to a gap;

(f) defining one or more mean field equations wherein said plurality of binary assignment variables are replaced by a plurality of continuous mean field variables, whereby said each mean field variable has a value between 0 and 1, and a plurality of forces that are proportional to said atomic distances squared;

(g) formulating an energy function, wherein:

said energy function includes a first cost for each said atomic distance wherein said distance is a weighted body transformation using said continuous mean field variables of said first protein while keeping said second protein fixed;

said energy function includes a second cost λ for each said gap by either said first protein or said second protein,

said energy function includes a third cost δ for a position-independent consecutive said gap;

said energy function includes a fourth cost for enforcing a constraint to satisfy that each said atom of said first protein either aligns with said atom of said second protein or to said gap;

(h) minimizing by an iterative process of said energy function and updating said continuous mean field variables in said mean field equations for a decreasing set of temperatures T until convergence to a predefined convergence value is reached; and

(i) after convergence rounding off said continuous mean field variables to either 0 or 1.

2. The method as set forth in claim 1, wherein said step of formulating an energy function further comprises a fifth cost for discouraging crossed matches.

3. The method as set forth in claim 1, wherein said second cost λ is a value between 0.01 and 0.5.

4. The method as set forth in claim 3, wherein said second cost λ for a α-site in a α-helix has a larger said second cost λ by a factor between 0.01 and 0.5 of said second cost λ.

5. The method as set forth in claim 3, wherein said second cost λ for a β-sheet has a larger said second cost λ by a factor between 0.01 and 0.5 of said second cost λ.

6. The method as set forth in claim 1, wherein said third cost δ is a function of said second cost λ divided by a value between 1 and 20.

7. The method as set forth in claim 1, wherein fourth cost includes a parameter γ with a value between 0 and 0.2.

8. The method as set forth in claim 1, wherein said step of minimizing by an iterative process includes an iteration parameter ε with a value between 0.5 and 0.95.

9. The method as set forth in claim 1, wherein said step of minimizing by an iterative process further comprises the step of initiating said temperature to a value between 1 and 100.

10. A method of using a mean field approach for protein structure alignment, comprising the steps of:

(a) providing a first protein with N_1 atoms;

(b) providing a second protein with N_2 atoms;

(c) making an initial alignment of said atoms of said first protein to said atoms of said second protein;

(d) calculating all atomic distances between said coordinates of said atoms of said first protein and said atomic coordinates of said atoms of said second protein;

(e) defining a matrix with a plurality of binary assignment variables wherein each binary assignment variable corresponding to either a match or to a gap;

(f) defining one or more mean field equations wherein said plurality of binary assignment variables are replaced by a plurality of continuous mean field variables, whereby said each mean field variable has a value between 0 and 1, and a plurality of forces that are proportional to said atomic distances squared;

(g) formulating an energy function, wherein:

said energy function includes a first cost for each said atomic distance wherein said distance is a weighted body transformation using said continuous mean field variables of said first protein while keeping said second protein fixed;

said energy function includes a second cost λ for each said gap by either said first protein or said second protein,

said energy function includes a third cost δ for a position-independent consecutive said gap;

said energy function includes a fourth cost for enforcing a constraint to satisfy that each said atom of said first protein either aligns with said atom of said second protein or to said gap;

(h) minimizing by an iterative process of said energy function and updating said continuous mean field variables in said mean field equations for a decreasing set of temperatures T until convergence to a predefined convergence value is reached; and

(i) after convergence rounding off said continuous mean field variables to either 0 or 1.

11. The method as set forth in claim 10, wherein said step of formulating an energy function further comprises a fifth cost for discouraging crossed matches.

12. The method as set forth in claim 10, wherein said second cost λ is a value between 0.01 and 0.5.

13. The method as set forth in claim 12, wherein said second cost λ for a α-site in a α-helix has a larger said second cost λ by a factor between 0.01 and 0.5 of said second cost λ.

14. The method as set forth in claim 12, wherein said second cost λ for a β-sheet has a larger said second cost λ by a factor between 0.01 and 0.5 of said second cost λ.

15. The method as set forth in claim 10, wherein said third cost δ is a function of said second cost λ divided by a value between 1 and 20.

16. The method as set forth in claim 10, wherein fourth cost includes a parameter γ with a value between 0 and 0.2.

17. The method as set forth in claim 10, wherein said step of minimizing by an iterative process includes an iteration parameter ε with a value between 0.5 and 0.95.

18. The method as set forth in claim 10, wherein said step of minimizing by an iterative process further comprises the step of initiating said temperature to a value between 1 and 100.